Seeing the Data When You Can't See the Data:
A Tour of Documentation of the CDW

October 27, 2014

By Margaret Gonsoulin, PhD

VIReC
RESEARCHERS' GUIDE TO VA DATA

Thanks to:

- Richard Pham, BISL/CDW
- Mark Dean, BISL/CDW
- Colleagues at VIReC

Definitions

- Corporate Data Warehouse (CDW) – national level database housing clinical, administrative and financial Veterans Health Administration information.

- Business Intelligence Service Line (BISL) – the group that architects the CDW.

- Metadata – documentation describing the data contents.

- VistA – Veterans Health Information Systems and Technology Architecture; an information system that has many applications, including the Electronic Health Record.

In today's talk…

- Continue the discussion that began in my last cyberseminar "First Time Research Users' Guide to CDW: Getting Started with this Relational Database"
- Identify the sites that hold documentation of CDW
- Review the content of materials in each site
- Learn how to use these resources

Poll #1: Your CDW Experience

Rate your level of experience with CDW data on a scale of 1 to 5…

- 1 Not worked with it at all

- 2

- 3

- 4

- 5 Very experienced with CDW

Poll #2: Have you seen "Getting Started with this Relational Database"?

Did you view our CDW cyberseminar, "First Time Research Users' Guide to CDW: Getting Started with this Relational Database" (live broadcast July 2014)?

- Yes
- No
- Not sure

Reminder – 2 Domain Types

- **Production Domains – contains tables that have been structured by BISL to support their re-joining.**
- **Raw Domains – contains tables that are direct extracts from the source system (e.g., VistA) that are simply being housed with no editing performed on them.**
- **These types of domains also vary in terms of their documentation.**

Example #1 -
A Production Domain

Let's say that you are doing research that pertains to allergic reactions. You do not currently have access to the data, but you want to learn about which measures exist in the CDW that are related to substances associated with allergic reactions. You decide to search the metadata for the production domain called Allergy.

5 Documentation Sites for Production Domains

We will explore 5 main sites housing information related to our example domain:

1. VHA Data Portal
2. VA Information Resource Center (VIReC)
3. CDW SharePoint
4. Data Architecture Repository (DAR)
5. VA Informatics and Computing Infrastructure (VINCI)
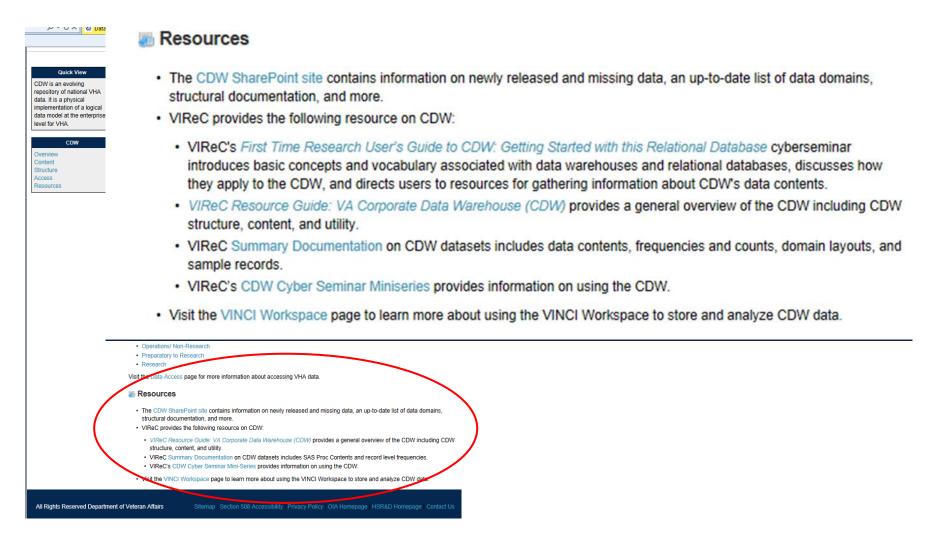
VHA Data Portal

- **This site provides:**
  - Links to all of the other sites mentioned in this talk
  - Information about how to apply for access to CDW data

VHA Data Portal

Direct Link- http://vaww.vhadataportal.med.va.gov/DataSources/CDW.aspx

VIReC's CDW Documentation Page

- **This consists of a set of products that allow a glimpse into CDW data before you have been granted official access.**
  - ▫ Lists of domains and tables
  - ▫ Formats of each column
  - ▫ Examples of records
  - ▫ Counts and frequencies

VIReC INTRANET

VA INFORMATION RESOURCE CENTER (VIReC)

VIReC Home
VA/CMS Home
About Us
New Users of VA Data
Data Transition to CDW
News & Updates
FAQs
Acronyms
HelpDesk

CDW Documentation

Overview

VIReC documentation for the Corporate Data Warehouse (CDW) includes domain layouts, data contents, sample records, and frequencies. This summary information is available by domain or schema.

Domain Layout

A list of schemas and tables for each CDW domain.

View Domain Layout

List of domains & tables

Data Contents

A list of data elements for each CDW domain.

View Data Contents

List of columns & their formats

Sample Records

Randomly selected records for each CDW domains.

View Sample Records

20 sample records

Frequencies & Counts

Record counts, null counts, and selected discrete value frequencies by schema.

View Frequencies & Counts

Frequencies for discrete values & counts of null values

http://vaww.virec.research.va.gov/CDW/Documentation.htm

VIReC
RESEARCHERS' GUIDE TO VA DATA
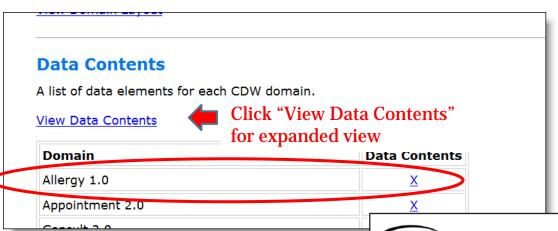
Domain Layout

A list of schemas and tables for each CDW domain.

View Domain Layout

VHA Corporate Data Warehouse (CDW) Domain Layout, June 2014

Domain	Schema	Table Name
Allergy 1.0	Allergy	AllergicReaction
	Allergy	AllergicReactionAssessment
	Allergy	Allergy
	Allergy	AllergyComment
	Allergy	AllergyDrugClass
	Allergy	AllergyDrugIngredient
	Dim	Reactant
	Dim	ReactantDrugClass
	Dim	ReactantDrugIngredient
	Dim	ReactantSynonym
	Dim	Reaction
	Dim	ReactionEffectiveDate
	Dim	ReactionSynonym
Appointment 2.0	Appt	Appointment
	Appt	AppointmentMultiple
	Dim	CancellationReason
Consult 2.0	Con	Consult
	Con	ConsultActivity

VIReC
RESEARCHERS' GUIDE TO VA DATA

View Domain Layout

Data Contents

A list of data elements for each CDW domain.

View Data Contents ⬅ **Click "View Data Contents" for expanded view**

Domain	Data Contents
Allergy 1.0	X
Appointment 2.0	X
Consult 2.0	

VIReC VHA Corporate Data Warehouse (CDW) Allergy 1.0 Domain Data Conten

About This Workbook

This workbook provides data contents for the Allergy 1.0 domain data in the Corporate Data Warehouse (CDW).

Select a table name from the list below to access a spreadsheet with its data contents. Table descriptions and fie the MetaData Report on CDW's SharePoint site, https://vaww.dwh.cdw.portal.va.gov/metadata/default.aspx.

Schema	Table
Allergy	AllergicReaction
Allergy	AllergicReactionAssessment
Allergy	Allergy
Allergy	AllergyComment
Allergy	AllergyDrugClass
Allergy	AllergyDrugIngredient
Dim	Reactant
Dim	ReactantDrugClass
Dim	ReactantDrugIngredient
Dim	ReactantSynonym
Dim	Reaction
Dim	ReactionEffectiveDate
Dim	ReactionSynonym

Spreadsheet Column Definitions

Field Name		the column name in each CDW table
	Data Type	Defines the type of value a field can contain.
	Length	For a character string or unicode data type, length represents the number of characters
SQL		represents the number of bytes that are used to store the number.
	Precision	the number of digits in a number excluding decimal point
	Scale	the number of digits to the right of the decimal point in a number

Data Content Dim.Reactant

VHA Corporate Data Warehouse (CDW) Allergy 1.0 Domain Data Contents, March 2014						
Table Name: Reactant						
Field Name	**SQL**				**SAS**	
	Data Type	**Length**	**Precision**	**Scale**	**Data Type**	**Length**
MasterEntryForVUIDFlag	CHAR	1			CHARACTER	1
NationalFlag	CHAR	1			CHARACTER	1
Reactant	VARCHAR	50			CHARACTER	50
ReactantIEN	VARCHAR	50			CHARACTER	50
ReactantSID	INT	4	10		NUMERIC	8
ReactantType	VARCHAR	50			CHARACTER	50
Sta3n	SMALLINT	2	5		NUMERIC	8
VUID	VARCHAR	50			CHARACTER	50
Go to About Sheet						

Sample Records

Randomly selected records for eac

View Sample Records

Domain	Sample Records
Allergy 1.0	X
Appointment 2.0	X
Consult 2.0	X

VIReC **VHA Corporate Data Warehouse (CDW) Allergy 1.0 Domain Sample Records**

About This Workbook

This workbook provides 20 randomly selected records from each table of the Allergy 1.0 domain in the Corporate Data W
records containing Protected Health Information (PHI) have been de-identified. Thus, while each record shows a comple
data fields, all PHI have been replaced by Xs for numeric data fields and YYYY-MM-DD or YYYY-MM-DD HH:MM:SS for

Select a table name from the list below to access a spreadsheet with its sample records. Table descriptions and field de
the MetaData Report on CDW's SharePoint site, https://vaww.dwh.cdw.portal.va.gov/metadata/default.aspx.

Schema	Table
Allergy	AllergicReaction
Allergy	AllergicReactionAssessment
Allergy	Allergy
Allgery	AllergyComment
Allergy	AllergyDrugClass
Allergy	AllergyDrugIngredient
Dim	Reactant
Dim	ReactantDrugClass
Dim	ReactantDrugIngredient
Dim	ReactantSynonym
Dim	Reaction
Dim	ReactionEffectiveDate
Dim	ReactionSynonym

20 sample records from Dim.Reactant

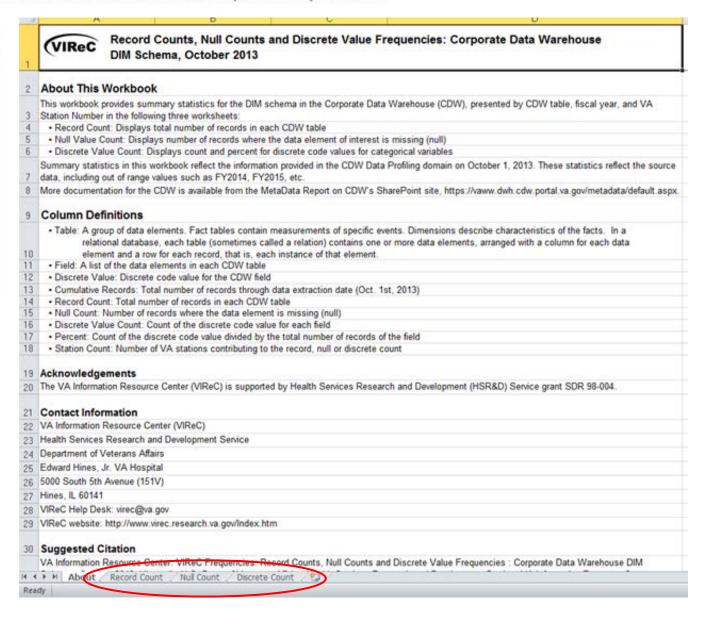| ReactantSID	ReactantIEN	Sta3n	Reactant	ReactantType	NationalFlag	MasterEntryForVUIDFlag	VUID
			VHA Corporate Data Warehouse (CDW) Allergy 1.0 Domain Sample Records, March 2014				
			Table Name: Dim.Reactant				
1200002951	689	590	WINE	DF	Y	Y	4636870
800006724	204	501	RADISHES	F	NULL	N	4636901
800005164	78	692	DIPHTHERIA TOXOID	DF	Y	Y	4538540
1000000991	274	695	CHIVES	F	Y	Y	4637377
1400002480	447	632	INDOOR MOLD	O	Y	Y	4690929
1200031173	755	581	PORTUGUESE MAN OF WAR STING	O	NULL	Y	4706313
1000006763	526	618	BUMBLE BEE STINGS	O	Y	Y	4636660
1000020594	208	502	BEETS	F	Y	Y	4636705
1200012281	813	558	PUDDING	F	NULL	Y	4538909
800010873	771	644	CAPERS	F	NULL	Y	4706335
1400003047	477	632	LUX SOAP	O	Y	Y	4636760
800001560	205	691	ZZZPHENOBARBITAL 30MG TAB	D	NULL	Y	4539166
800011487	246	653	ADHESIVES	O	Y	Y	4637425
1000013651	603	695	SALT	F	Y	Y	4636878
1400000409	406	608	ELASTOPLAST TAPE	O	Y	Y	4636627
1000000240	150	676	CRAB	F	Y	Y	4636668
1000018954	502	549	PALM OIL	F	Y	Y	4636723
1400009360	243	518	CLEANING SOLUTIONS	O	Y	Y	4637371
1000008888	644	564	TEGADERM	O	Y	Y	4636809
1000016933	520	556	ONIONS	DF	Y	Y	4637334

Go to About Sheet

Free text

Discrete

Frequencies & Counts

Record counts, null counts, and selected discrete value frequencies by schema.

View Frequencies & Counts

Schema	Frequency
ALLERGY	X
APPT	X
CHEM	X
CON	X
CPRSORDER	X
DENTAL	X
DIM	X

(VIReC) Record Counts, Null Counts and Discrete Value Frequencies: Corporate Data Warehouse DIM Schema, October 2013

About This Workbook

This workbook provides summary statistics for the DIM schema in the Corporate Data Warehouse (CDW), presented by CDW table, fiscal year, and VA Station Number in the following three worksheets:

- Record Count: Displays total number of records in each CDW table
- Null Value Count: Displays number of records where the data element of interest is missing (null)
- Discrete Value Count: Displays count and percent for discrete code values for categorical variables

Summary statistics in this workbook reflect the information provided in the CDW Data Profiling domain on October 1, 2013. These statistics reflect the source data, including out of range values such as FY2014, FY2015, etc.

More documentation for the CDW is available from the MetaData Report on CDW's SharePoint site, https://vaww.dwh.cdw.portal.va.gov/metadata/default.aspx.

Column Definitions

- Table: A group of data elements. Fact tables contain measurements of specific events. Dimensions describe characteristics of the facts. In a relational database, each table (sometimes called a relation) contains one or more data elements, arranged with a column for each data element and a row for each record, that is, each instance of that element.
- Field: A list of the data elements in each CDW table
- Discrete Value: Discrete code value for the CDW field
- Cumulative Records: Total number of records through data extraction date (Oct. 1st, 2013)
- Record Count: Total number of records in each CDW table
- Null Count: Number of records where the data element is missing (null)
- Discrete Value Count: Count of the discrete code value for each field
- Percent: Count of the discrete code value divided by the total number of records of the field
- Station Count: Number of VA stations contributing to the record, null or discrete count

Acknowledgements

The VA Information Resource Center (VIReC) is supported by Health Services Research and Development (HSR&D) Service grant SDR 98-004.

Contact Information

VA Information Resource Center (VIReC)
Health Services Research and Development Service
Department of Veterans Affairs
Edward Hines, Jr. VA Hospital
5000 South 5th Avenue (151V)
Hines, IL 60141
VIReC Help Desk: virec@va.gov
VIReC website: http://www.virec.research.va.gov/Index.htm

Suggested Citation

VA Information Resource Center. VIReC Frequencies: Record Counts, Null Counts and Discrete Value Frequencies : Corporate Data Warehouse DIM

About / Record Count / Null Count / Discrete Count

Ready

Record Count

	A	B	C
1	**SCHEMA: DIM**		
2	**Table**	**Cumulative Records**	
3		**Record Count**	**Station Count**
118	REACTANT	106,386	138
119	RELEASEEVENT	6,669	132
120	RELIGION	10,993	138

Null Count

	A	B	C	D
1	**SCHEMA: DIM**			
2	**Table** ▼	**Field** ▼	**Cumulative Records**	
3			**Null Count**	**Station Count**
1120	REACTANT	ALLERGYTERMIEN	0	138
1121	REACTANT	ALLERGYTYPE	218	138
1122	REACTANT	MASTERENTRYFORVUIDFLAG	0	138
1123	REACTANT	NATIONALALLERGYFLAG	14,092	138
1124	REACTANT	VUID	0	138
1125	RELEASEEVENT	COPYACTIVEORDERSFLAG	2,412	132
1126	RELEASEEVENT	DISPLAYTEXT	974	132
1127	RELEASEEVENT	INACTIVATEDDATETIME	4,647	132
1128	RELEASEEVENT	INSTITUTIONIEN	1,200	132

Discrete Value Counts

	A	B	C	D	E	F
1	**SCHEMA: DIM**					
2	**Table** ▼	**Field** ▼	**Discrete Value** ▼	**Cumulative Records**		
3				**Discrete Value Count**	**Percent**	**Station Count**
145669	REACTANT	MASTERENTRYFORVUIDFLAG	*	8	0.01	8
145670	REACTANT	MASTERENTRYFORVUIDFLAG	N	1,330	1.25	129
145671	REACTANT	MASTERENTRYFORVUIDFLAG	Y	105,048	98.74	130
145672	REACTANT	NATIONALALLERGYFLAG		14,092	13.25	130
145673	REACTANT	NATIONALALLERGYFLAG	*	8	0.01	8
145674	REACTANT	NATIONALALLERGYFLAG	N	92,286	86.75	130
145675	RELEASEEVENT	TYPEOFEVENT		221	4.03	16
145676	RELEASEEVENT	TYPEOFEVENT	*Missing*	4	0.07	4
145677	RELEASEEVENT	TYPEOFEVENT	*Unknown at this time*	4	0.07	4

CDW SharePoint

- Here a user can find information about:
  - Table contents
  - VHA data origins
  - Linking keys
  - Data formats

https://vaww.dwh.cdw.portal.va.gov/Pages/CDWHome.aspx

Metadata for Production Domains

https://vaww.dwh.cdw.portal.va.gov/metadata/default.aspx

Entity Relationship Diagrams

Entity Relationship Diagrams

- **Blue boxes represent fact tables & pink represent dimension tables.**
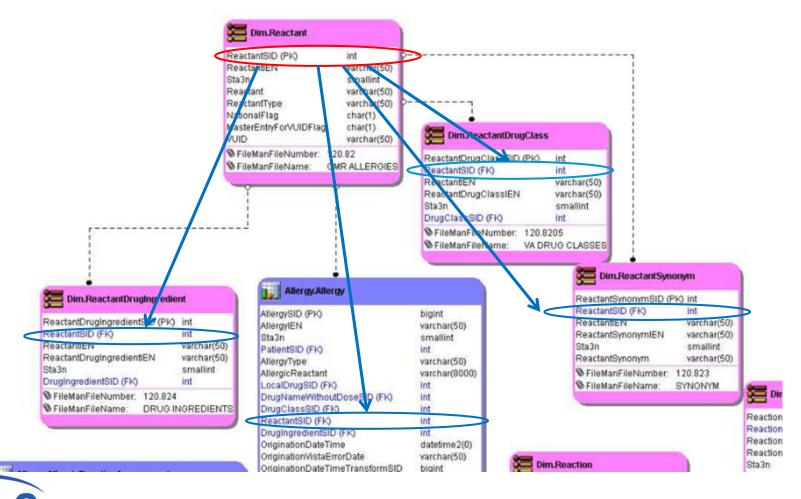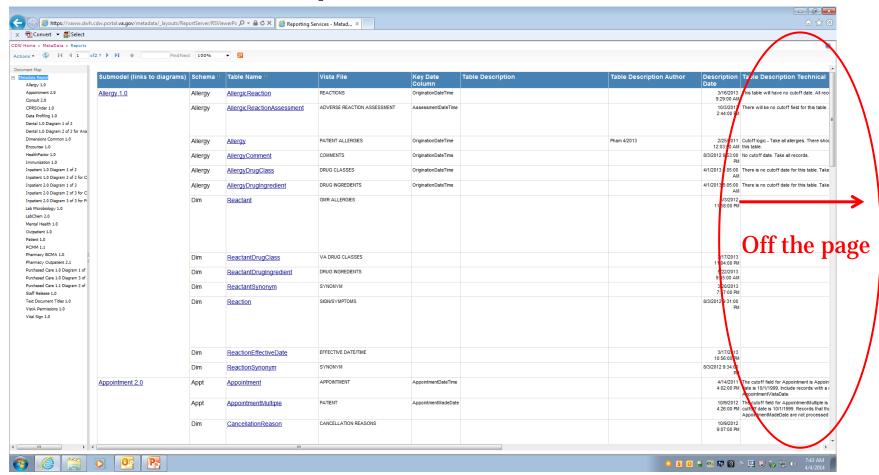- **The lines represent links between tables.**

ER Diagram (magnified)

Table Descriptions

Off the page

Descriptions of each table as a whole

Descriptions of Each Table (off to side)

Table Descriptions (edited in excel)

	Submodel (links to diagrams)	Schema	Table Name	Vista File	Vista File Description
3	Allergy 1.0	Allergy	AllergicReaction	REACTIONS	These are the reactions observed for this allergy/adverse reaction
4		Allergy	AllergicReactionAssessment	ADVERSE REACTION ASSESSMENT	This file is a listing of all the patients who have been asked aboutallergies/adverse reactions. It contains a pointer to File 2 (PATIENT), aflag to indicate if the patient has or does not have a Allergy/AdverseReaction, the person making the assessment a the date/time of theassessment.
5		Allergy	Allergy	PATIENT ALLERGIES	Contains patient allergy/adverse reaction information.
6		Allergy	AllergyComment	COMMENTS	List of comments entered about this reaction.
7		Allergy	AllergyDrugClass	DRUG CLASSES	List of VA drug classes to which patient had reaction for thisallergy/adverse reaction.
8		Allergy	AllergyDrugIngredient	DRUG INGREDIENTS	List of drug ingredients to which patient had reaction for thisallergy/adverse reaction.
9		Dim	Reactant	GMR ALLERGIES	Contains a listing of allergies from which user can select. Per VHA directive XXX, this file has been "locked down" by DataStandardization (DS). The file definition (i.e. data dictionary shallnot be modified. All additions, changes and deletions to entries in thefile shall be done by Enterprise Reference Terminology (ERT) using theMaster File Server (MFS), provided by Common Services (CS). Creatingand/or editing locally defin fields in the file are not permitted. Useof locally defined fields th were created prior to VHA Directive XXXshall not be supported.
10		Dim	ReactantDrugClass	VA DRUG CLASSES	List of VA Drug classes that comprise this reactant.
11		Dim	ReactantDrugIngredient	DRUG INGREDIENTS	List of drug ingredients that comprise this particular allergy.

VIReC Description --- Dim.Reactant contains the name of the reactant (e.g., "mustard," "insulin" or "epoxy") along with the type of reactant (e.g., food or drug). In addition, it makes note of whether or not the allergen or reactant is distributed through a national releaser or is locally added.
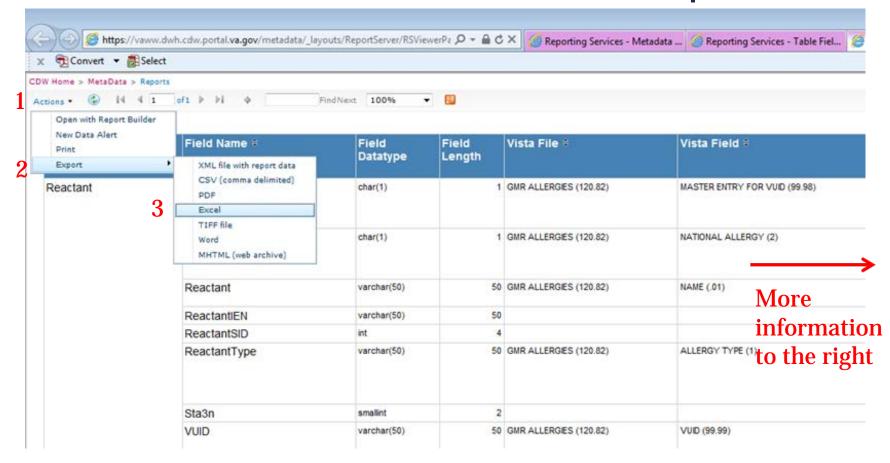
Metadata Reports for Each Table

Metadata reports for each table allows users to see descriptions for each column

CDW Table-level Metadata Report

10/27/2014

Dim.Reactant Metadata (edited in excel)

Table Name	Field Name	Field Datatype	Field Length	Vista File	Vista Field	Primary Key	Foreign Key	Vista Field Description
Reactant	MasterEntryForVUIDFlag	char(1)	1	GMR ALLERGIES (120.82)	MASTER ENTRY FOR VUID (99.98)	N	N	This field identifies the Master entry for a VUID associatedwith a Term/Concept.
	NationalFlag	char(1)	1	GMR ALLERGIES (120.82)	NATIONAL ALLERGY (2)	N	N	Indicates whether this allergy was distributed with the national releaseor is a locally added allergy.
	Reactant	varchar(50)	50	GMR ALLERGIES (120.82)	NAME (.01)	N	N	The name of the allergy/adverse reaction.
	ReactantIEN	varchar(50)	50			N	N	
	ReactantSID	int	4			Y	N	
	ReactantType	varchar(50)	50	GMR ALLERGIES (120.82)	ALLERGY TYPE (1)	N	N	This field contains the type(s) for this allergy/adverse reaction . Theuser can enter the type(s) separated by commas, or the following codes:D=Drug, F=Food, O=Other. If codes are used, do not use commas to separatemultiple codes. Examples of valid entries are: DRUG or DRUG, FOOD or Dor DF or OTHER.
	Sta3n	smallint	2			N	N	
	VUID	varchar(50)	50	GMR ALLERGIES (120.82)	VUID (99.99)	N	N	VHA Unique ID (VUID). A unique meaningless integerassigned to reference terms VHA wide.

This field contains the type(s) for this allergy/adverse reaction . Theuser can enter the type(s) separated by commas, or the following codes:D=Drug, F=Food, O=Other. If codes are used, do not use commas to separatemultiple codes. Examples of valid entries are: DRUG or DRUG, FOOD or Dor DF or OTHER.
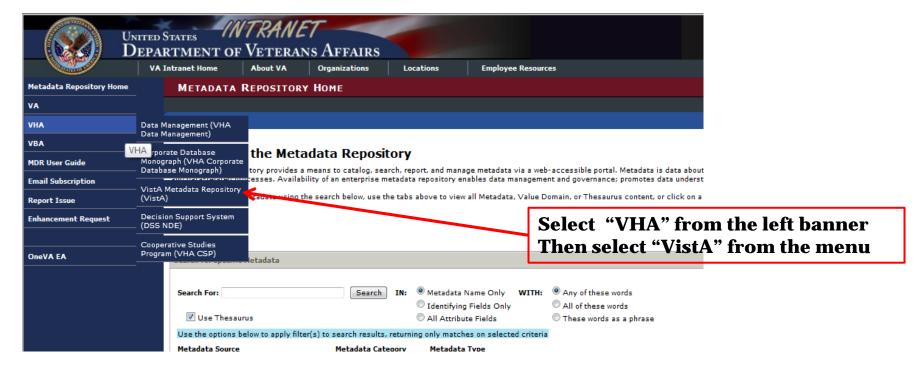
VIReC
RESEARCHERS' GUIDE TO VA DATA

Data Architecture Repository (DAR)

- This site serves as a repository for metadata from around the VA.

- It includes metadata for VHA information and VistA information.

- VistA is one of the primary sources of CDW data.

- Referencing these files can help clarify the meaning of CDW data.

DAR's VistA Metadata

Once inside, you will want to search by file name or number or field name or number. The files and fields are listed in the CDW Metadata.
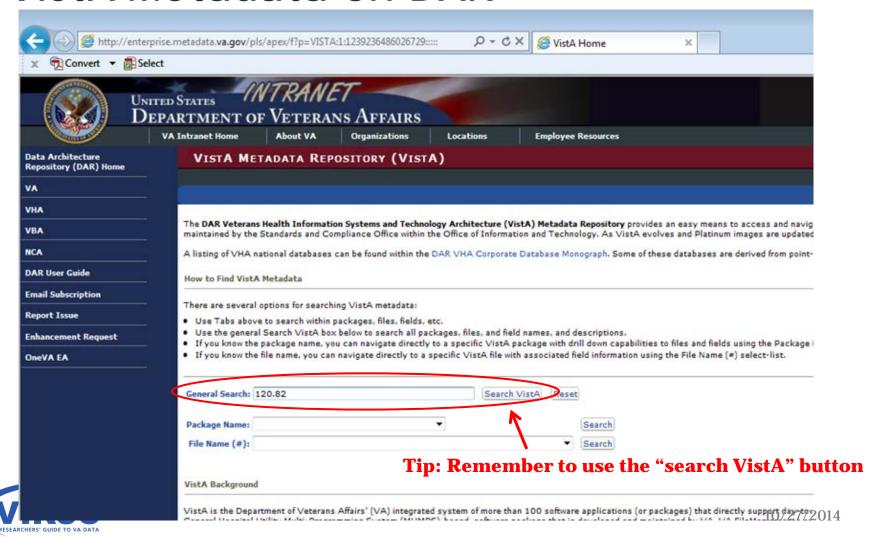
http://vaausdarmul81/pls/apex/f?p=2000:1:246212839773488

Dim.Reactant Metadata (edited in excel)

Table Name	Field Name	Field Datatype	Field Length	Vista File	Vista Field	Primary Key	Foreign Key	Vista Field Description
Reactant	MasterEntryForVUIDFlag	char(1)	1	GMR ALLERGIES (120.82)	MASTER ENTRY FOR VUID (99.98)	N	N	This field identifies the Master entry for a VUID associated with a Term/Concept.
	NationalFlag	char(1)	1	GMR ALLERGIES (120.82)	NATIONAL ALLERGY (2)	N	N	Indicates whether this allergy was distributed with the national release or is a locally added allergy.
	Reactant	varchar(50)	50	GMR ALLERGIES (120.82)	NAME (.01)	N	N	The name of the allergy/adverse reaction.
	ReactantIEN	varchar(50)	50			N	N	
	ReactantSID	int	4			Y	N	
	ReactantType	varchar(50)	50	GMR ALLERGIES (120.82)	ALLERGY TYPE (1)	N	N	This field contains the type(s) for this allergy/adverse reaction . The user can enter the type(s) separated by commas, or the following codes:D=Drug, F=Food, O=Other. If codes are used, do not use commas to separate multiple codes. Examples of valid entries are: DRUG or DRUG, FOOD or D or DF or OTHER.
	Sta3n	smallint	2			N	N	
	VUID	varchar(50)	50	GMR ALLERGIES (120.82)	VUID (99.99)	N	N	VHA Unique ID (VUID). A unique meaningless integer assigned to reference terms VHA wide.

The file of interest is "GMR Allergies" #120.82

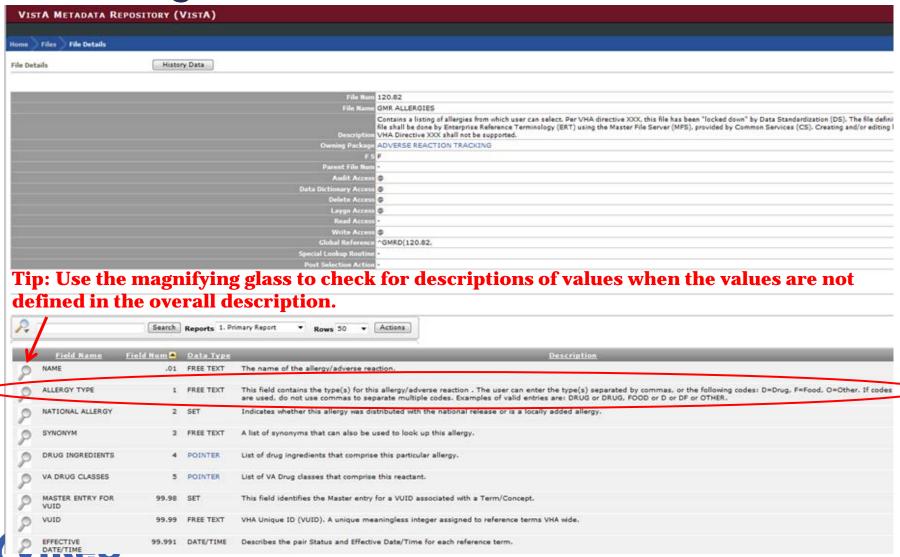**The field of interest is "Allergy Type" #1**

VistA Metadata on DAR
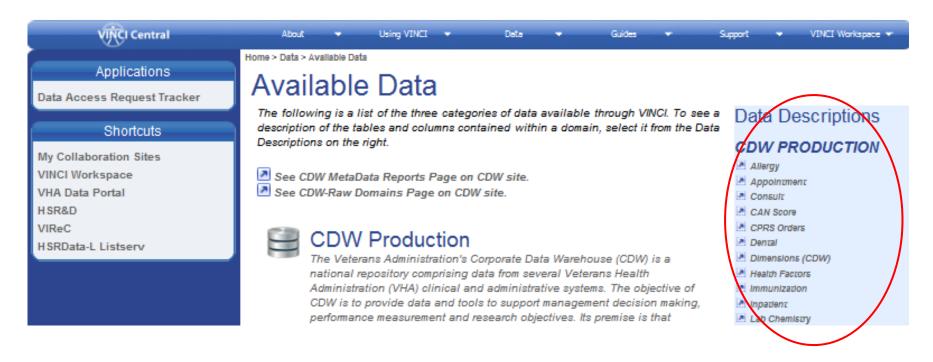
Finding the file 120.82- DAR

Looking at the fields in 120.82

VISTA METADATA REPOSITORY (VISTA)

Home > Files > File Details

File Details   History Data

File Num	120.82
File Name	GMR ALLERGIES
Description	Contains a listing of allergies from which user can select. Per VHA directive XXX, this file has been "locked down" by Data Standardization (DS). The file defini file shall be done by Enterprise Reference Terminology (ERT) using the Master File Server (MFS), provided by Common Services (CS). Creating and/or editing VHA Directive XXX shall not be supported.
Owning Package	ADVERSE REACTION TRACKING
F S	F
Parent File Num	-
Audit Access	@
Data Dictionary Access	@
Delete Access	@
Laygo Access	@
Read Access	-
Write Access	@
Global Reference	^GMRD(120.82,
Special Lookup Routine	-
Post Selection Action	-

Tip: Use the magnifying glass to check for descriptions of values when the values are not defined in the overall description.

Search Reports 1. Primary Report ▼ Rows 50 ▼ Actions

	Field Name	Field Num	Data Type	Description
🔍	NAME	.01	FREE TEXT	The name of the allergy/adverse reaction.
🔍	ALLERGY TYPE	1	FREE TEXT	This field contains the type(s) for this allergy/adverse reaction . The user can enter the type(s) separated by commas, or the following codes: D=Drug, F=Food, O=Other. If codes are used, do not use commas to separate multiple codes. Examples of valid entries are: DRUG or DRUG, FOOD or D or DF or OTHER.
🔍	NATIONAL ALLERGY	2	SET	Indicates whether this allergy was distributed with the national release or is a locally added allergy.
🔍	SYNONYM	3	FREE TEXT	A list of synonyms that can also be used to look up this allergy.
🔍	DRUG INGREDIENTS	4	POINTER	List of drug ingredients that comprise this particular allergy.
🔍	VA DRUG CLASSES	5	POINTER	List of VA Drug classes that comprise this reactant.
🔍	MASTER ENTRY FOR VUID	99.98	SET	This field identifies the Master entry for a VUID associated with a Term/Concept.
🔍	VUID	99.99	FREE TEXT	VHA Unique ID (VUID). A unique meaningless integer assigned to reference terms VHA wide.
🔍	EFFECTIVE DATE/TIME	99.991	DATE/TIME	Describes the pair Status and Effective Date/Time for each reference term.

RESEARCHERS' GUIDE TO VA DATA

VINCI's Documentation

Here one will find additional documentation related to various CDW domains. The documentation is organized by domain. It contains some descriptions of meaning of content found in tables and columns. Documentation related to the Allergy domain is currently in transition.

http://vaww.vinci.med.va.gov/vincicentral/Data.html

Poll #3: About You

Which of the following best describes your role in the VA?

❑ Research Investigator / PI

❑ Data Manager / Analyst

❑ Project Coordinator

❑ Operations

❑ Other

Next example… A Raw Domain

Now, you are doing research that requires access to X-ray results. You do not currently have access to the data, but you want to learn about which measures exist in the CDW that are related to radiology. You decide to search the metadata for the raw domain called Radiology.

VIReC
RESEARCHERS' GUIDE TO VA DATA

3 Documentation Sites for Raw Domains

- We will explore 3 main sites housing information related to our example domain:
  1. CDW SharePoint
  2. Data Architecture Repository (DAR)
  3. VA Informatics and Computing Infrastructure (VINCI)

A Raw Domain – Radiology Example

CDW Raw Reports

Find the names of the raw domain here.
Scroll down to find radiology.

Radiology Domain Report

| Radiology v3 Schema: Radiology since Mar 2011 | Hamid.Saoudian@va.gov; Evan.Lott@va.gov; Mark.Love@va.gov; Joseph.Eccleston@va.gov; Ted.Fletcher@va.gov; Jack.Bates@va.gov; Kathleen.Frisbee@va.gov; Shawn.Hardenbrook@va.gov; Ron.Moody@va.gov; Jason.Ray@va.gov; Timothy.Dove@va.gov; Christa.Kramer@va.gov; Cheryl.Canody@va.gov; Rebecca.Gandillon@va.gov; Scot.Dingman@va.gov; Jody.Pasell@va.gov; James.Prelerson@va.gov; | Twice per Month ——— 2nd & 4th Friday | SEP 04 2014 | SEP 21 2014 | 72 Hours | 2 B recs 415 Gb | Basic (CDW_Full) | Initial requirements came from VISN20 ER diagrams and VINCI team. POC: Hamid Saoudian Radiology skipped 02/17/2012 |

Click here

To arrive at this report

CDW-Raw Radiology Extracts – v3
3/19/2012 3:42 PM

CDW-Raw Radiology Extracts – v3
  Summary
    Key Contacts
    Extraction Schedule
    Extracted Tables
    Change History
    Permissions
    Current Inventory
  Best Practices
    Sample Queries
    Date Handling
    Indexes
  Extraction Metadata
    Performance Tuning
    Radiology Grouping
    Extraction Summary
    Grouping with Tables
    SQL Extract Queries
    Extraction Errors
    VistA Files & Fields
  VistA Site numbers

Summary

This document contains information about the CDW Raw Extraction of the Radiology grouping. This domain was a priority domain by the CDW Governance Board. This extract will be from all 130 VistA sites, starting

The data can be accessed using the views in the CDWWork database on vhacdwa06.

Key Contacts

Current Information as of 01/30/2012

Project /Domain	Radiology
Current Stakeholders	Evan.Lott@va.gov Nsta.Yao@va.gov Mark.Love@va.gov Joseph.Eccleston@va.gov Nstall.Anderson@va.gov Ted.Fletcher@va.gov Jack.Bates@va.gov Kathleen.Frisbee@va.gov
Current Point of Contact	Hamid.Saoudian@va.gov

Extraction Schedule

Radiology is extracted according to the CDW-Raw Weekly Extraction Schedule, currently the first and third Fridays of the month.

Extracted Tables

The original requirements came from the VINCI team on 3/16/2011. The list included all of the files on the VISN 20 data model (below) and some additional tables needed by VINCI.

Change History

12/9/2011 – v3

Added these for Kathy Frisbee
rad_modality_defined_term_73_1
rdlgycptbyprodrtype_73_2

10/27/2014

List of Tables

8/29/2011

#	Schema	Table	Count	KBy
1	Radiology	complication_types_78_1	1,151	
2	Radiology	contractsharing_agreements_34	574	
3	Radiology	diagnostic_codes_78_3	5,887	
4	Radiology	examination_status_72	8,015	
5	Radiology	imaging_locations_79_1	3,932	
6	Radiology	imaging_stop_codes_71_5	1,761	
7	Radiology	imaging_type_79_2	1,211	
8	Radiology	nuc_med_exam_data_70_2	929,887	
9	Radiology	nuc_x_radiopharmaceutica_70_21	1,106,965	
10	Radiology	procedure_modifiers_71_2	8,416	
11	Radiology	procedux_type_of_imaging_71_23	11,253	
12	Radiology	radnuc_med_common_procedu_71_3	30,029	
13	Radiology	radnuc_med_division_79	502	
14	Radiology	radnuc_med_orders_75_1	114,631,527	
15	Radiology	radnuc_med_patient_70	13,003,262	
16	Radiology	radnuc_med_procedures_71	199,197	
17	Radiology	radnuc_med_reason_75_2	6,956	
18	Radiology	radnuc_med_reports_74	127,459,417	1
19	Radiology	radnuc_med_rx_other_case_74_05	12,464,589	
20	Radiology	radnuc_med_x_descendents_71_05	45,890	
21	Radiology	radnucx_contrast_media_71_0125	37,276	
22	Radiology	radnucx_registered_exams_70_02	118,788,661	
23	Radiology	radnux_hold_description_75_125	358,560	
24	Radiology	radnux_imaging_locations_79_01	3,279	
25	Radiology	radnx_clinical_history_f_75_11	355,277,404	

- Scroll down to find a list of tables in the domain.

- Notice the names are using the VistA naming traditions.

- The first table in the list is Radiology.complication_types_78_1

- That corresponds to VistA file "Complication Types" #78.1

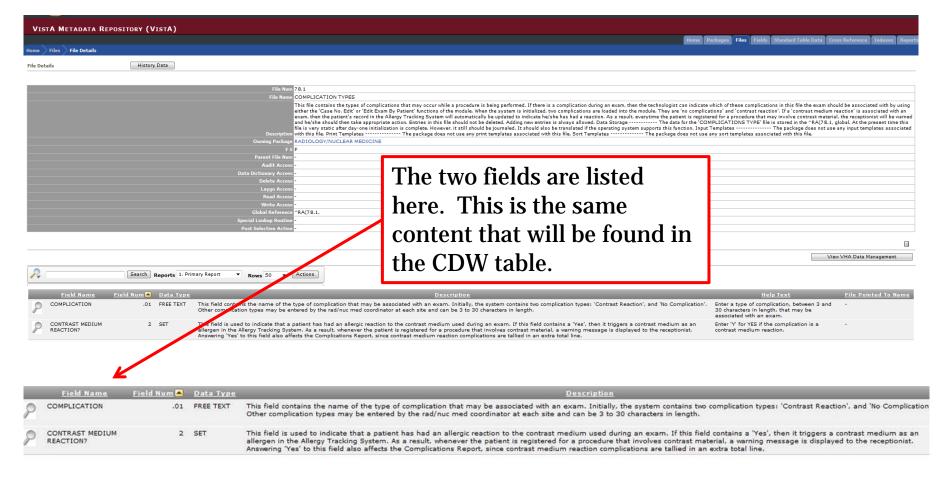# VistA Documentation in DAR #78.1
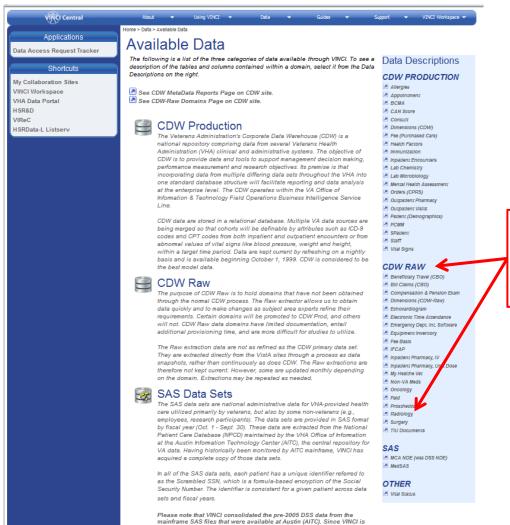

Search for file 78.1

Select "complication types"

Examine Content

The two fields are listed here. This is the same content that will be found in the CDW table.

VINCI Raw Documentation

Also find reports about raw domains in the VINCI data documentation.

VINCI Radiology Report

VINCI Data Description
Raw Radiology Extract

Contents

10/27/2014

VINCI Table Description

Table Name	Table description
complication_types_78_1	This file contains the types of complications that may occur while a procedure is being performed. If there is a complication during an exam, then the technologist can indicate which of these complications in this file the exam should be associated with by using either the 'Case No. Edit' or 'Edit Exam By Patient' functions of the module. When the system is initialized, two complications are loaded into the module. They are 'no complications' and 'contrast reaction'. If a 'contrast medium reaction' is associated with an exam, then the patient's record in the Allergy Tracking System will automatically be updated to indicate he/she has had a reaction. As a result, everytime the patient is registered for a procedure that may involve contrast material, the receptionist will be warned and he/she should then take appropriate action. Entries in this file should not be deleted. Adding new entries is always allowed.

VINCI Field Descriptions

complication_types_78_1

ColumnName	DataType	Description
sta3n	Smallint	
ComplicationTypesIEN	Varchar(50)	
Complication	Varchar(50)	This field contains the name of the type of complication that may be associated with an exam. Initially, the system contains two complication types: 'Contrast Reaction', and 'No Complication'. Other complication types may be entered by the rad/nuc med coordinator at each site and can be 3 to 30 characters in length.
ContrastMediumReaction	Varchar(30)	Field Codes : Y:YES;N:NO; This field is used to indicate that a patient has had an allergic reaction to the contrast medium used during an exam. If this field contains a 'Yes', then it triggers a contrast medium as an allergen in the Allergy Tracking System. As a result, whenever the patient is registered for a procedure that involves contrast material, a warning message is displayed to the receptionist. Answering 'Yes' to this field also affects the Complications Report, since contrast medium reaction complications are tallied in an extra total line.

Limitation of these sources

- **The main limitation of these documentation sources is the frequent lack value descriptions.**
- **4 places to look:**
    - DAR (see tip on slide 38 – magnifying glass)
    - VIReC Sample Records
    - VIReC Discrete Counts (most tables)
    - VIReC Factbooks (Consult Domain available now)

Summary/Conclusion

- Some variety of these sites will prove useful in learning about the CDW content in which you are interested.

- These primary variation are based on whether or not the data is in a production domain or in a raw domain.

- However, there is also endless variety of considerations regarding where and when you find the specific documentation you are seeking.

Questions?

Margaret Gonsoulin, PhD

VIReC@va.gov

708-202-2413

Content Specific Resources

Additional Documentation & Assistance (1)

Domain	Source - Title
Consult	VIReC Factbook- Corporate Data Warehouse (CDW) Consult 2.0 Domain
Inpatient	CDW Insights Days Presentation 3/26/ 14 & Data Quality Report - Guide to Inpatient 2.0 Contents CDW Insights Days Presentation 3/26/14- Data Transformation from VistA files to CDW
Lab Microbiology	VIReC Researcher's Notebook - Issue 1: How do I identify the causative organism in CDW Lab Microbiology? An Example Using Staphylococcus Aureus
Lab Chem	VIReC Cyberseminar- A Practical Guide to Using the Corporate Data Warehouse LabChem Data: VA Informatics and Computing
Mental Health	CDW Insights Days Presentation 3/26/14- Tele-Mental Health: Preliminary Analysis of Unique Patients
Outpatient	VIReC Cyberseminar - Identification of Outpatient Workload in CDW Data Quality Report - Applying NPCD Business Rules to CDW Outpatient Visits to calculate Workload
Patient	Data Quality Report - Work on Race and Ethnicity Data Quality Report - Identifying Unique Patients in CDW (SPatient Only) Data Quality Report - Identifying Veterans in the CDW Data Quality Report - Identifying Test Patients in CDW
Pharmacy	CDW Insights Days Presentation 4/30/14- Applied Research in the VINCI Environment
Vital Status	Data Quality Report - CDW Date of Death Analysis and Vital Status File Comparison

Additional Documentation & Assistance (2)

- Data Quality Page:
  http://vaww.vhadataquality.va.gov/index.php?lang=en

- CDW Insights Days:
  https://vaww.dwh.cdw.portal.va.gov/Support/Training/SitePages/Home.aspx

- VIReC Cyberseminars:
  http://www.virec.research.va.gov/Resources/Cyberseminars.asp#Upcoming

- VIReC HelpDesk:
  http://vaww.virec.research.va.gov/Support/HelpDesk.htm

- HSRData-L ListServ:
  http://vaww.virec.research.va.gov/Support/HSRData-L.htm

Next session: January 26, 2015

Getting the Information You Need from CDW: SQL Starter Language

Margaret Gonsoulin, PhD